

画像の複比を検索キーに用いた文書検索

平松 正行[†] 江崎 信行[‡] 竹下 鉄夫[‡]

豊田工業高等専門学校 専攻科情報科学専攻[†] 豊田工業高等専門学校 情報工学科[‡]

1. はじめに

Apple の iPad, Amazon.com の kindle, そして iPhone に代表されるスマートフォンの普及により, 日本でも近い将来に電子書籍, 電子新聞のような電子文書が普及すると考えられる。

紙などに印刷された文書と電子文書の大きな違いは, 視認性, 省スペース性, 検索性などが挙げられる。本研究では, このうちの検索性を考える。印刷された文書と異なり, 電子文書はタイトルやその内容, また作成日時などをキーとして PC などで検索することができる。

内容を検索する場合, その文書に含まれている文字列を検索キーとして行うのが一般的である。しかし, 文書をスキャナなどにより OCR などの処理をせずに単にコンピュータに取り込んだ場合や検索利用者が精通していない言語の検索を行う場合, 文字列をキーとする検索が難しい。

このような問題を解決する手法として, デジタルカメラなどで撮影した画像の特徴量を検索キーとして同一の文書を検索する手法が考案されている[1]。

新聞はスクラップして保存することが多く, その出典がわからなくなることが多いため, 本稿では, 電子文書としてスキャンした新聞を扱う。検索方法として「複比の投票に基づく」方法を用いて, 新聞の切り抜きなどの新聞の一部からその出典を検索する新聞検索システムの作成について述べる。

2. 新聞検索システム

2.1. システム概要

本稿で作成するシステムは, スキャナを用いて取り込んだ新聞画像を, デジタルカメラやウェブカメラで撮影した検索キー画像(新聞の切り抜きなどの新聞の一部)を用いて検索することを目的とする。システムの概要図を図 1 に示す。

スキャナにより取り込む新聞画像は, 射影歪みがなくノイズがほとんどない画像である。一方, 検索キーとなる画像は, 新聞の一部で, 射影歪みがある画像が含まれる(図 2)。

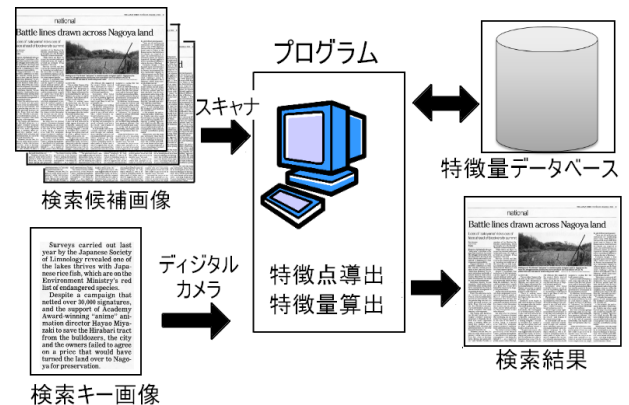


図 1: 開発システム概要図



図 2: 検索キー画像の例
(左: 正対画像, 右: 射影歪みがある画像)

また, 新聞は分かち書きのされる言語で書かれているもの(英語, ドイツ語など)を対象とする。

2.2. データベースへの登録

2.2.1. 新聞画像の取り込み

新聞の取り込みはスキャナで行う。家庭などで普及している A4 スキャナでは 1 度で全てスキャンすることができなかったため, 新聞 1 面を 4 回にわけてスキャンした。今回は 4 つに分けた画像データは連結したりせずそのまま扱う。

2.2.2. 複比の導出

データベースへ登録するのは新聞の画像ではなく, 新聞の特徴点の配置から得られた特徴量である複比の値である。

複比の導出は文献[1]に従い, 2 値化, 単語領域の切り出し, ラベリング, 重心導出により特徴点を求めた後, それぞれの特徴点の近傍 5 点ごとに時計回りに並べ替えた後, 以下の式を用いて複比を計算する。

Document Search Using Cross-ratio of Picture

[†] HIRAMATSU, Masayuki

[‡] ESAKI, Nobuyuki

[‡] TAKESHITA, Tetsuo

Computer Science Advanced Course, Toyota National College of Technology ([†])

Information & Computer Engineering, Toyota National College of Technology ([‡])

$$\rho(P_1, P_2, P_3, P_4, P_5) = \frac{S(P_1, P_2, P_3)S(P_1, P_4, P_5)}{S(P_1, P_2, P_4)S(P_1, P_3, P_5)}$$

ここで、 $S(P_1, P_2, P_3)$ は点 P_1, P_2, P_3 を頂点とする三角形の面積である。検索キーとなる画像が射影歪みを受けている場合があるので、射影変換に不変な値である複比[2]を特徴量に用いている。どの点を始点にするかにより特徴点 1 点につき 5 通りの複比が求められる(近傍点の時計回りの順序が射影変換に不変)。ここで求めた複比を不等間隔な 10 段階に離散化してデータベースへ登録する。

2.3. データベースからの検索

2.3.1. 検索キー画像の取り込み

検索キーは、デジタルカメラやウェブカメラで取り込んだ画像を対象としている。また、検索キーは射影歪みが含まれる。

2.3.2. 複比の導出

候補画像と同様に 2.2.2.節の方法で、検索キー画像に対して特徴点とその特徴量である複比の導出を行う。複比は 5 点の複比の順によって変化するので、順序を入れ替え 5 通りの複比を求める。

2.3.3. 検索

2.3.2.節で求めた複比をデータベースへ問い合わせる。5 つの複比が一致している特徴点があれば、その文書に投票する。これにより多数の投票を得た文書が検索結果候補となる。

2.4. 検索結果の表示

検索結果は、GUI で作成した画面上に 1 番投票された新聞の縮小画像および推定場所を表示する。

3. 実験

3.1. 実験方法

システムの評価は、新聞 THE DAILY YOMIURI No.21653(2011/01/05)のある 1 面のデータを 2.2 節の方法を用いてデータベースに登録し、検索に登録した新聞の一部を検索キー行う。この実験により検索キーの場所を当てることができたかを確かめる。

次に、新聞全体(22 面)をスキャンし、同様にデータベースに登録し、検索キーとして新聞画像の一部を検索キーとして入力し検索精度について調べる。

3.2. 実験結果

特徴点の数は 1 面あたり約 10000 個であり、データベースの作成には 1 面あたり約 3 分、新聞 1 面で 1 時間程度の時間を要した(データベース作成環境は CPU: Pentium M 1.73GHz, メモリ: 2GB である)。特徴点の数は平均すると 1 面あたり約 10000 個だったが、テレビ欄、株式欄など 2 万点以上の特徴点を持つものもあった。

まず、新聞のうち、1 面、12 面、22 面をデータベースとして使い、それぞれスキャナで取得した 5 つの検索キーで正しく場所を当てることができるかを調べた。この検索キーには射影歪みがない。その結果、15 個の検索キーのうち、10 個で正しく場所を当てることができた。検索時には 30 秒程度の時間がかかった(検索実行環境は CPU: Core2Quad 3.0GHz, メモリ: 4GB である)。

一方、新聞全体からの検索についても実験を行った。この実験では、スキャナ、デジタルカメラで撮影した正対画像、射影歪みのある画像を入力したが、いずれも検索は成功しなかった。

4. 考察

1 面のデータの場所を当てる実験では、検索データベースの数を限定することで、検索に成功することができた。失敗の 5 例では、検索キーの特徴点が多分でなかったことが原因であると考えられる。ゆえに、検索キーは大きいほうが検索が成功する可能性が高くなると推測される。

新聞全体からの検索が成功しなかった理由として、実験結果で述べたように株式欄やテレビ欄は他のページと比べて極端に特徴点数が多いため、検索キーより求めた複比が、テレビ欄や株式欄より算出された複比と偶然一致してしまったため、正しい結果よりも得票数が多くなってしまったためと考えられる。

5. まとめと今後の課題

本稿では、複比の投票に基づく文書画像検索と射影歪み補正[1]の方によって、新聞の切り抜きなど新聞の一部からその出典を調べるシステムの提案を行った。

今後、新聞検索システムに改良を加え、精度を向上させていくことが課題である。この際、新聞に含まれている図の情報を活用しようと考えている。最終的に、OCR ソフトを利用した文字認識による新聞の検索と本システムでの新聞の検索について比較検証していく。

参考文献

- [1] 中居友弘・黄瀬浩一・岩村雅一、複比の投票に基づく文書画像検索と射影歪み補正、画像の認識・理解シンポジウム(MIRU2005), 2005
- [2] Tomas Suk and Jan Flusser, Point-based projective invariants, Pattern Recognition 33, 2000